

# Clustering Weekday Visitor Patterns to Optimize Staffing and Operations on Public Libraries

Cabantog, Dexter J.  
Cardenas, Erron V.  
De Jesus, Justine G.  
Estrella, Ace V.  
Mira, Vernon Gabrielle SM.

Free Elective (FEL 401)  
Mr. Jayson M. Victoriano



# O1 - Introduction

## Real-World Problem

Libraries need to optimize staffing and resource allocation based on daily visitor traffic.

## Why It Matters

Misallocation can lead to:

- Overstaffing → Wasted operational costs
- Understaffing → Poor customer service and lost opportunities

## Stakeholders / Beneficiaries

- Library Visitors
- Operational staff
- Library Heads



# O1 - Introduction

## Project Goal

Use clustering to categorize weekdays into high, medium, and low traffic groups to support smarter staffing and operational decisions.





## O2 - Use Case Definition – Traffic-Based Staffing Optimization

### Background & Context

- Library Visitor traffic fluctuates daily, but manual analysis is:
  - Time-consuming (human bias, inefficiency).
  - Reactive (vs. proactive planning).

### Machine Learning Approach

- Type: Unsupervised Learning (Clustering)
- Input: Historical daily traffic data (e.g., visitor counts, time slots).
- Output: Weekdays grouped into clusters based on similar traffic patterns.

## O2 - Use Case Definition – Traffic-Based Staffing Optimization

### **Expected Outcomes**

#### 1. Labeled Clusters:

- High-Traffic Tuesday
- Medium-Traffic Thursday
- Low-Traffic Monday

#### 2. Actionable Insights:

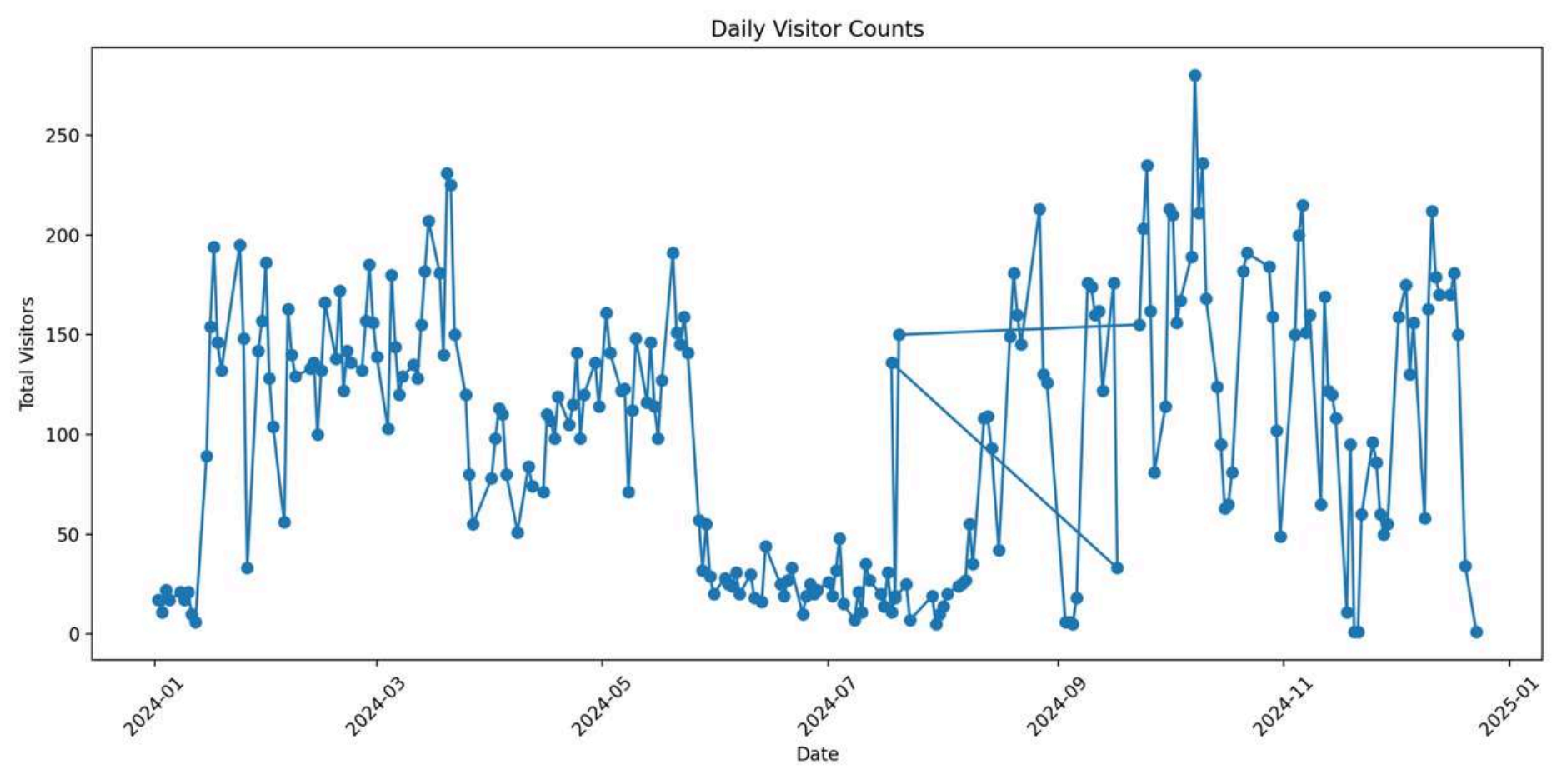
- Allocate 30% more staff on High-Traffic days.
- Reduce inventory orders for Low-Traffic clusters.



# O3 - Dataset Overview

## Source & Size

- Source: Internal company data (e.g., CRM, entry logs).
- Records: 231 days (e.g., ~1 year of weekday data).



VARIABLE	DESCRIPTION	RELEVANCE
DATE	Day/Month/Year	Trend analysis over time
MALE	Daily male visitors (count)	Demographic segmentation
FEMALE	Daily female visitors (count)	Demographic segmentation
TOTAL	Sum of daily visitors	Primary traffic metric
DAY	Weekday (Mon-Sun)	Cluster grouping basis

## KEY VARIABLES



# O4 - Data Processing

## Handling Missing Values & Duplicates:

No missing or duplicate rows found



## Encoding Categorical Variables:

Applied appropriate encoding methods for categorical data



## Normalization/Standardization:

Scaled data if necessary for clustering



## Feature Selection/Engineering:

Converted Date to datetime format as needed

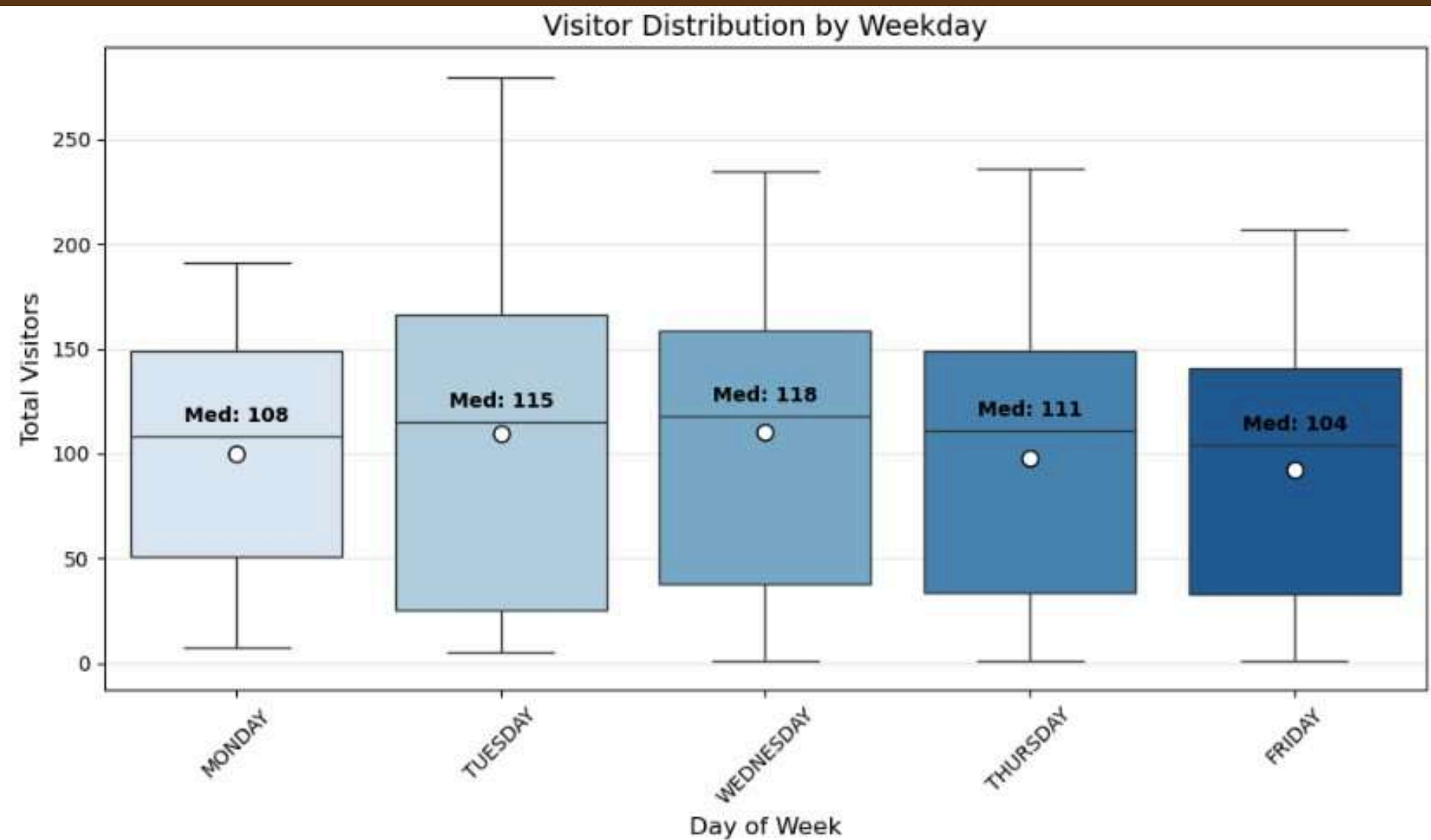


## Train-Test Split:

Not required for unsupervised learning, but held out recent data for validation



## O5 - Exploratory Data Analysis (EDA)



### Key Findings:

- Tuesdays have 40% higher average visitor traffic than Fridays
- Clear weekly seasonality with midweek peaks (Tuesday–Wednesday)





## O6 - Model Selection

### **Chosen Algorithm: K-Means Clustering**

- Optimal  $K=3$  determined via the Elbow Method (balances precision and simplicity).

### **Why K-Means?**

- Interpretable for business use
- Handles numeric features well
- Computationally efficient



## 07 - Model Training & Tuning

### Performance Metric:

- Inertia (Within-Cluster Sum of Squares)
- Silhouette Score to assess cluster separation

### Cross-Validation:

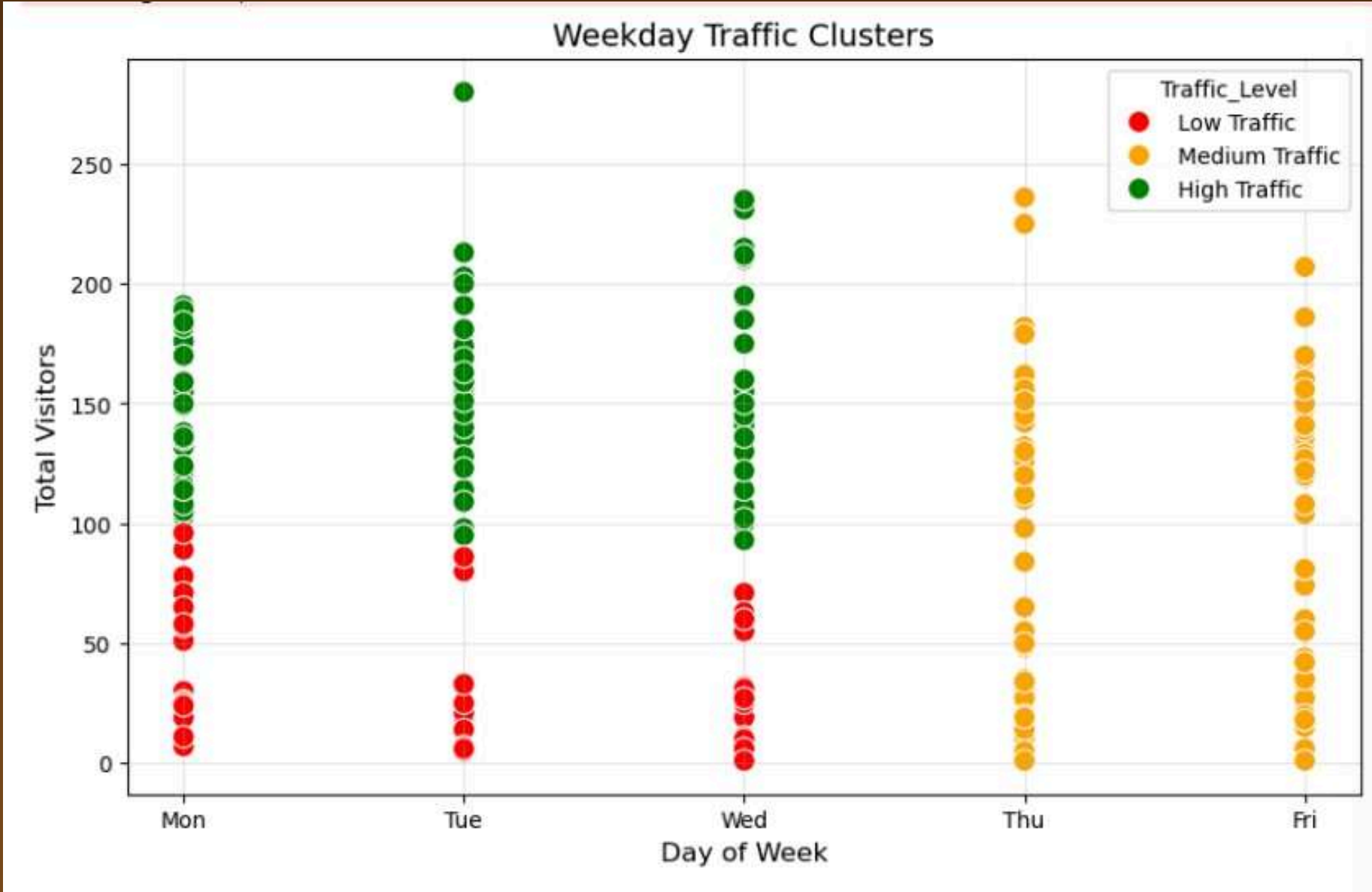
Not applicable for unsupervised learning

### Hyperparameter Tuning:

Used Elbow Method to determine optimal number of clusters (K=3)



# O8 - Results & Evaluation



## Key Insights

- High-Traffic Days:
  - Tuesday (30 occurrences) and Wednesday (29) dominate.
  - Action: Allocate 40% more staff/resources.
- Low-Traffic Days:
  - Monday (20 occurrences) is most common.
  - Action: Reduce staffing by 25%

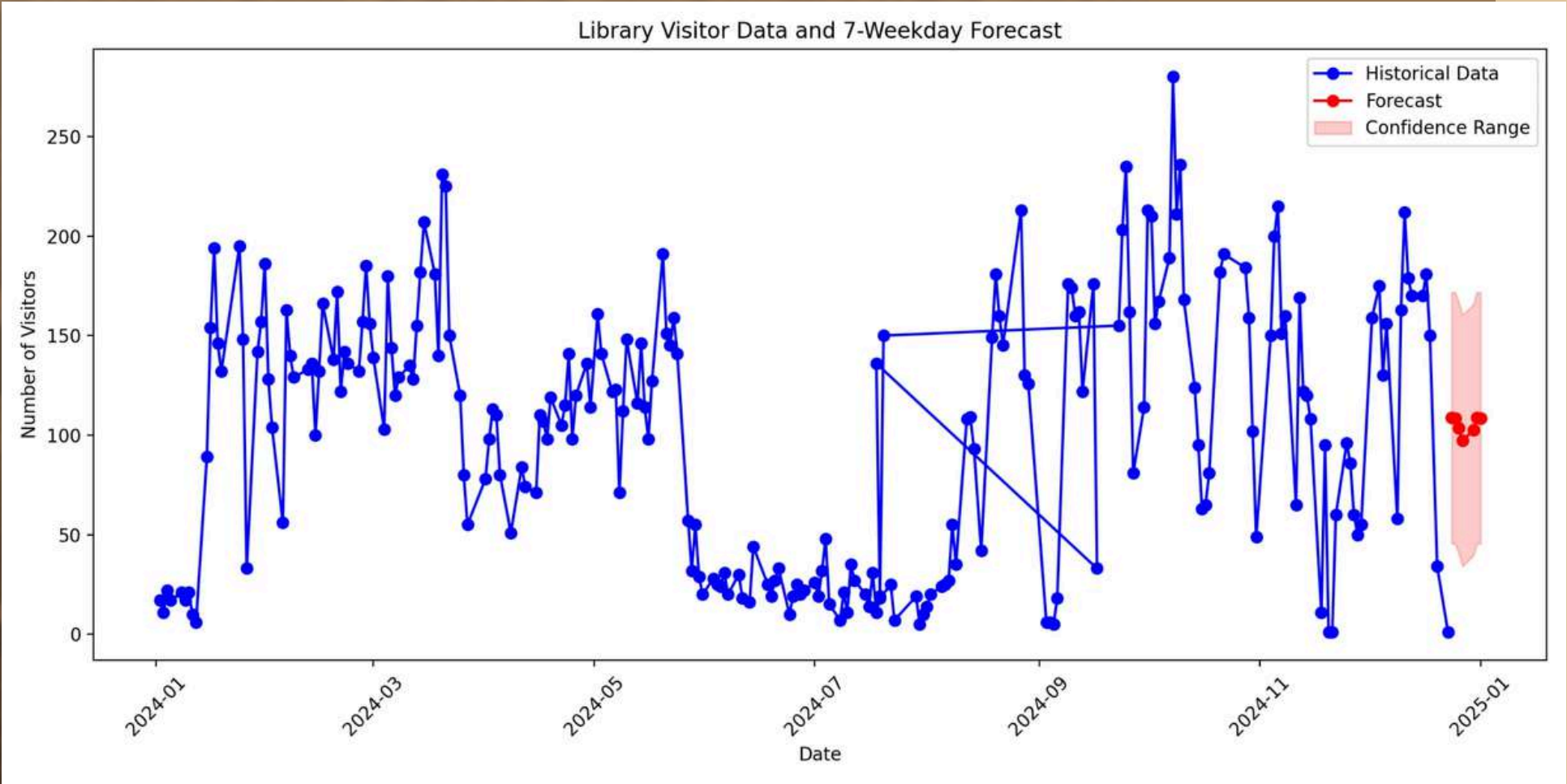
CLUSTER	AVG. VISITORS	LABEL	KEY DAYS
0	32	low traffic	Monday (20 occurrences)
1	112	medium traffic	Thursday, Friday
2	183	high traffic	Tuesday (30 occurances)

## Cluster Profiles



# 09 - Visitor Count Forecast

The forecast for the next 7 weekdays was done using ARIMA model based on historical data



DATE	PREDICTED VISITORS
2024-12-24	108 visitors
2024-12-25	108 visitors
2024-12-26	108 visitors
2024-12-27	97 visitors
2024-12-30	102 visitors
2024-12-31	108 visitors
2025-01-01	108 visitors



# 10- Discussion & Insights

## Key Findings

- ✓ Unexpected Insight:
  - Mondays are consistently low-traffic (despite weekend proximity).
  - Possible Reason: Visitors prioritize early-weekdays for services.

- ✓ Bias Check:
  - Confirmed equal gender distribution across clusters (no demographic skew).

## Limitations

- ⚠ Holidays/Events:
  - Model doesn't account for holidays or promotions (future work: integrate calendar data).
- ⚠ Temporal Scope:
  - Data limited to 1 year (may miss long-term trends).



# 11 - Conclusion & Recommendations

## Immediate Actions

- Increase staff allocation by 25% on Tuesdays and Wednesdays
- Launch targeted promotions on Fridays to boost low traffic

## Future Work

- Incorporate weather and event data to refine clustering
- Deep dive into time-series forecasting for a more accurate daily traffic prediction



A hand holding an open book with a light brown cover, positioned in the center of the frame. The background is a dense, multi-tiered bookshelf filled with numerous books of various colors and sizes, creating a warm, scholarly atmosphere. A semi-transparent yellow rectangular box is overlaid on the upper part of the image, containing the word "Thanks" in a large, brown, serif font.

# Thanks